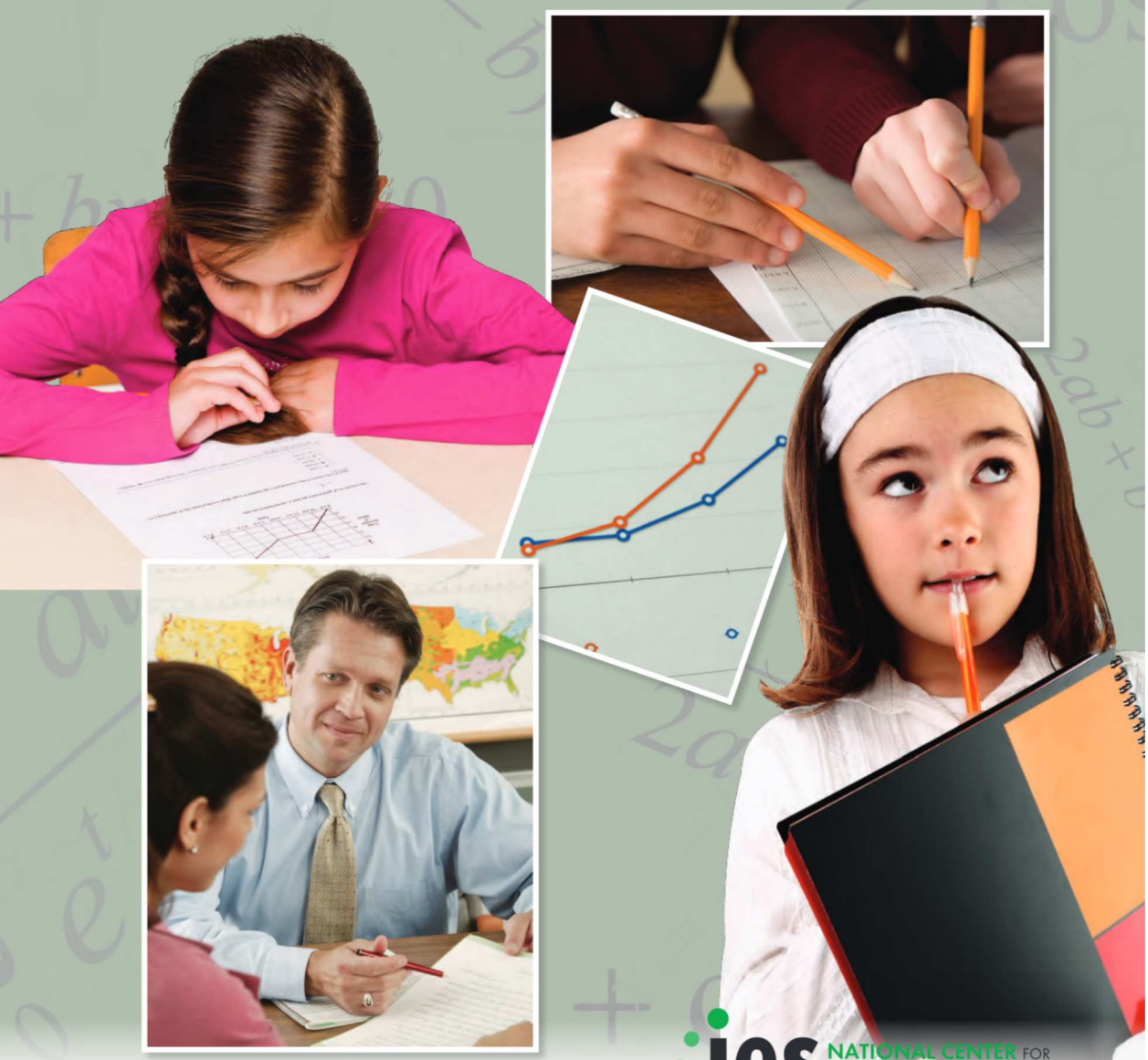
Using Student Achievement Data to Support Instructional Decision Making



NCEE 2009-4067
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

The Institute of Education Sciences (IES) publishes practice guides in education to bring the best available evidence and expertise to bear on the types of challenges that cannot currently be addressed by a single intervention or program. Authors of practice guides seldom conduct the types of systematic literature searches that are the backbone of a meta-analysis, although they take advantage of such work when it is already published. Instead, authors use their expertise to identify the most important research with respect to their recommendations and conduct a search of recent publications to ensure that the research supporting the recommendations is up-to-date.

Unique to IES-sponsored practice guides is that they are subjected to rigorous external peer review through the same office that is responsible for independent reviews of other IES publications. A critical task for peer reviewers of a practice guide is to determine whether the evidence cited in support of particular recommendations is up-to-date and that studies of similar or better quality that point in a different direction have not been ignored. Because practice guides depend on the expertise of their authors and their group decision making, the content of a practice guide is not and should not be viewed as a set of recommendations that in every case depends on and flows inevitably from scientific research.

The goal of this practice guide is to formulate specific and coherent evidence-based recommendations for use by educators and education administrators to create the organizational conditions necessary to make decisions using student achievement data in classrooms, schools, and districts. The guide provides practical, clear information on critical topics related to data-based decision making and is based on the best available evidence as judged by the panel. Recommendations presented in this guide should not be construed to imply that no further research is warranted on the effectiveness of particular strategies for data-based decision making.

Using Student Achievement Data to Support Instructional Decision Making

September 2009

Panel

Laura Hamilton (Chair)
RAND Corporation

Richard Halverson
University of Wisconsin–Madison

Sharnell S. Jackson
Chicago Public Schools

Ellen Mandinach
CNA Education

Jonathan A. Supovitz
University of Pennsylvania

Jeffrey C. Wayman
The University of Texas at Austin

Staff

Cassandra Pickens
Emily Sama Martin
Mathematica Policy Research

Jennifer L. Steele
RAND Corporation

NCEE 2009-4067
U.S. DEPARTMENT OF EDUCATION

ies NATIONAL CENTER FOR EDUCATION EVALUATION AND REGIONAL ASSISTANCE
Institute of Education Sciences

This report was prepared for the National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, under Contract ED-07-CO-0062 by the What Works Clearinghouse, operated by Mathematica Policy Research.

Disclaimer

The opinions and positions expressed in this practice guide are the authors' and do not necessarily represent the opinions and positions of the Institute of Education Sciences or the U.S. Department of Education. This practice guide should be reviewed and applied according to the specific needs of the educators and education agency using it, and with the full realization that it represents the judgments of the review panel regarding what constitutes sensible practice, based on the research available at the time of publication. This practice guide should be used as a tool to assist in decision making rather than as a "cookbook." Any references within the document to specific education products are illustrative and do not imply endorsement of these products to the exclusion of other products that are not referenced.

U.S. Department of Education
Arne Duncan
Secretary

Institute of Education Sciences
John Q. Easton
Director

National Center for Education Evaluation and Regional Assistance
John Q. Easton
Acting Commissioner

September 2009

Hamilton, L., Halverson, R., Jackson, S., Mandinach, E., Supovitz, J., & Wayman, J. (2009). Using student achievement data to support instructional decision making (NCEE 2009-4067). Washington, DC: National Center for Education Evaluation and Regional Assistance, Institute of Education Sciences, U.S. Department of Education. Retrieved from http://ies.ed.gov/ncee/wwc/publications/practiceguides/.

What Works Clearinghouse Practice Guide citations begin with the panel chair, followed by the names of the panelists listed in alphabetical order.

This report is available on the IES website at http://ies.ed.gov/ncee and http://ies.ed.gov/ncee/wwc/publications/practiceguides/.

Alternative formats
On request, this publication can be made available in alternative formats, such as Braille, large print, audiotape, or computer diskette. For more information, call the Alternative Format Center at 202–205–8113.

Using Student Achievement Data to Support Instructional Decision Making

Contents

List of tables

List of figures

List of examples

Introduction

As educators face increasing pressure from federal, state, and local accountability policies to improve student achievement, the use of data has become more central to how many educators evaluate their practices and monitor students' academic progress.[1] Despite this trend, questions about how educators should use data to make instructional decisions remain mostly unanswered. In response, this guide provides a framework for using student achievement data to support instructional decision making. These decisions include, but are not limited to, how to adapt lessons or assignments in response to students' needs, alter classroom goals or objectives, or modify student-grouping arrangements. The guide also provides recommendations for creating the organizational and technological conditions that foster effective data use. Each recommendation describes action steps for implementation, as well as suggestions for addressing obstacles that may impede progress. In adopting this framework, educators will be best served by implementing the recommendations in this guide together rather than individually.

The recommendations reflect both the expertise of the panelists and the findings from several types of studies, including studies that use causal designs to examine the effectiveness of data use interventions, case studies of schools and districts that have made data-use a priority, and observations from other experts in the field. The research base for this guide was identified through a comprehensive search for studies evaluating academically oriented data-based decision-making interventions and practices. An initial search for literature related to data use to support instructional decision making in the past 20 years yielded more than 490 citations. Of these, 64 used experimental, quasi-experimental, and single subject designs to examine whether data use leads to increases in student achievement. Among the studies ultimately relevant to the panel's recommendations, only six meet the causal validity standards of the What Works Clearinghouse (WWC) and were related to the panel's recommendations.[2]

To indicate the strength of evidence supporting each recommendation, the panel relied on the WWC standards for determining levels of evidence, described below and in Table 1. It is important for the reader to remember that the level of evidence rating is not a judgment by the panel on how effective each of these recommended practices will be when implemented, nor is it a judgment of what prior research has to say about the effectiveness of these practices. The level of evidence ratings reflect the panel's judgment of the validity of the existing literature to support a causal claim that when these practices have been implemented in the past, positive effects on student academic outcomes were observed. They do not reflect judgments of the relative strength of these positive effects or the relative importance of the individual recommendations.

A strong rating refers to consistent and generalizable evidence that an intervention strategy or program improves outcomes.[3]

A moderate rating refers either to evidence from studies that allow strong causal conclusions but cannot be generalized with assurance to the population on which a recommendation is focused (perhaps because the findings have not been widely

1. Knapp et al. (2006).

2. Reviews of studies for this practice guide applied WWC Version 1.0 standards. See Version 1.0 standards at http://ies.ed.gov/ncee/wwc/pdf/wwc_version1_standards.pdf.

3. Following WWC guidelines, improved outcomes are indicated by either a positive, statistically significant effect or a positive, substantively important effect size (i.e., greater than 0.25).

replicated) or to evidence from studies that are generalizable but have more causal ambiguity than that offered by experimental designs (e.g., statistical models of correlational data or group comparison designs for which equivalence of the groups at pretest is uncertain).

A low rating refers to evidence either from studies such as case studies and descriptive studies that do not meet the standards for moderate or strong evidence or from expert opinion based on reasonable extrapolations from research and theory. A low level of evidence rating indicates that the panel did not identify a body of research demonstrating effects of implementing the recommended practice on student achievement. The lack of a body of valid evidence may simply mean that the recommended practices are not feasible or are difficult to study in a rigorous, experimental fashion.[4] In other cases, it means that researchers have not yet studied a practice or that there is weak or conflicting evidence of effectiveness. Policy interest in topics of current study thus can arise before a research base has accumulated on which recommendations can be based.

Under these circumstances, the panel examined the research it identified on the topic and combined findings from that research with its professional expertise and judgments to arrive at recommendations. However, that a recommendation has a low level of evidence should not be interpreted as indicating that the panel believes the recommendation is unimportant. The panel has decided that all five recommendations are important and, in fact, encourages educators to implement all of them to the extent that state and district resources and capacity allow.

4. For more information, see the WWC Frequently Asked Questions page for practice guides, http://ies.ed.gov/ncee/wwc/references/idocviewer/doc.aspx?docid=15&tocid=3.

Table 1. Institute of Education Sciences levels of evidence for practice guides

Strong	In general, characterization of the evidence for a recommendation as strong requires both studies with high internal validity (i.e., studies whose designs can support causal conclusions) and studies with high external validity (i.e., studies that in total include enough of the range of participants and settings on which the recommendation is focused to support the conclusion that the results can be generalized to those participants and settings). Strong evidence for this practice guide is operationalized as r □ A systematic review of research that generally meets WWC standards (see http://ies.ed.gov/ncee/wwc/) and supports the effectiveness of a program, practice, or approach with no contradictory evidence of similar quality; OR r □ Several well-designed, randomized controlled trials or well-designed quasi-experiments that generally meet WWC standards and support the effectiveness of a program, practice, or approach, with no contradictory evidence of similar quality; OR r □ One large, well-designed, randomized controlled, multisite trial that meets WWC standards and supports the effectiveness of a program, practice, or approach, with no contradictory evidence of similar quality; OR r □ For assessments, evidence of reliability and validity that meets the Standards for Educational and Psychological Testing.[a]
Moderate	In general, characterization of the evidence for a recommendation as moderate requires studies with high internal validity but moderate external validity or studies with high external validity but moderate internal validity. In other words, moderate evidence is derived from studies that support strong causal conclusions but generalization is uncertain or studies that support the generality of a relationship but the causality is uncertain. Moderate evidence for this practice guide is operationalized as r □ Experiments or quasi-experiments generally meeting WWC standards and supporting the effectiveness of a program, practice, or approach with small sample sizes and/or other conditions of implementation or analysis that limit generalizability and no contrary evidence; OR r □ Comparison group studies that do not demonstrate equivalence of groups at pretest and, therefore, do not meet WWC standards but that (1) consistently show enhanced outcomes for participants experiencing a particular program, practice, or approach and (2) have no major flaws related to internal validity other than lack of demonstrated equivalence at pretest (e.g., only one teacher or one class per condition, unequal amounts of instructional time, highly biased outcome measures); OR r □ Correlational research with strong statistical controls for selection bias and for discerning influence of endogenous factors and no contrary evidence; OR r □ For assessments, evidence of reliability that meets the Standards for Educational and Psychological Testing[b] but with evidence of validity from samples not adequately representative of the population on which the recommendation is focused.
Low	In general, characterization of the evidence for a recommendation as low means that the recommendation is based on expert opinion derived from strong findings or theories in related areas and/or expert opinion buttressed by direct evidence that does not rise to the moderate or strong level. Low evidence is operationalized as evidence not meeting the standards for the moderate or strong level.

a. American Educational Research Association, American Psychological Association, and National Council on Measurement in Education (1999).

b. Ibid.

The What Works Clearinghouse standards and their relevance to this guide

In terms of the levels of evidence indicated in Table 1, the panel relied on WWC evidence standards to assess the quality of evidence supporting educational programs and practices. The WWC evaluates evidence for the causal validity of instructional programs and practices according to WWC standards. Information about these standards is available at http://ies.ed.gov/ncee/wwc/pdf/wwc_version1_standards.pdf. The technical quality of each study is rated and placed into one of three categories:

- Meets Evidence Standards for randomized controlled trials and regression discontinuity studies that provide the strongest evidence of causal validity.
- Meets Evidence Standards with Reservations for all quasi-experimental studies with no design flaws and randomized controlled trials that have problems with randomization, attrition, or disruption.
- Does Not Meet Evidence Screens for studies that do not provide strong evidence of causal validity.

Following the recommendations and suggestions for carrying out the recommendations, Appendix D presents more information on the research evidence that supports each recommendation.

The panel would like to thank Cassandra Pickens, Emily Sama Martin, Dr. Jennifer L. Steele, and Mathematica and RAND staff members who participated in the panel meetings, characterized the research findings, and drafted the guide. We also appreciate the help of the many WWC reviewers who contributed their time and expertise to the review process, and Sarah Wissel for her support of the intricate logistics of the project. In addition, we would like to thank Scott Cody, Kristin Hallgren, Dr. Shannon Monahan, and Dr. Mark Dynarski for their oversight and guidance during the development of the practice guide.

Dr. Laura Hamilton
Dr. Richard Halverson
Ms. Sharnell S. Jackson, Ed.M.
Dr. Ellen Mandinach
Dr. Jonathan A. Supovitz
Dr. Jeffrey C. Wayman

Using Student Achievement Data to Support Instructional Decision Making

Overview

Recent changes in accountability and testing policies have provided educators with access to an abundance of student-level data, and the availability of such data has led many to want to strengthen the role of data for guiding instruction and improving student learning. The U.S. Department of Education recently echoed this desire, calling upon schools to use assessment data to respond to students' academic strengths and needs.[5] In addition, spurred in part by federal legislation and funding, states and districts are increasingly focused on building longitudinal data systems.[6]

Although accountability trends explain why more data are available in schools, the question of what to do with the data remains primarily unanswered. Data provide a way to assess what students are learning and the extent to which students are making progress toward goals. However, making sense of data requires concepts, theories, and interpretative frames of reference.[7] Using data systematically to ask questions and obtain insight about student progress is a logical way to monitor continuous improvement and tailor instruction to the needs of each student. Armed with data and the means to harness the information data can provide, educators can make instructional changes aimed at improving student achievement, such as:

- prioritizing instructional time;[8]
- targeting additional individual instruction for students who are struggling with particular topics;[9]
- more easily identifying individual students' strengths and instructional interventions that can help students continue to progress;[10]
- gauging the instructional effectiveness of classroom lessons;[11]
- refining instructional methods;[12] and
- examining schoolwide data to consider whether and how to adapt the curriculum based on information about students' strengths and weaknesses.[13]

5. American Recovery and Reinvestment Act of 2009; U.S. Department of Education (2009); Obama (2009).

6. Aarons (2009).

7. Knapp et al. (2006).

8. Brunner et al. (2005).

9. Brunner et al. (2005); Supovitz and Klein (2003); Wayman and Stringfield (2006).

10. Brunner et al. (2005); Forman (2007); Wayman and Stringfield (2006).

11. Halverson, Prichett, and Watson (2007); Supovitz and Klein (2003).

12. Halverson, Prichett, and Watson (2007); Fiarman (2007).

13. Marsh, Pane, and Hamilton (2006); Kerr et al. (2006).

Scope of the practice guide

The purpose of this practice guide is to help K–12 teachers and administrators use student achievement data to make instructional decisions intended to raise student achievement. The panel believes that the responsibility for effective data use lies with district leaders, school administrators, and classroom teachers and has crafted the recommendations accordingly.

This guide focuses on how schools can make use of common assessment data to improve teaching and learning. For the purpose of this guide, the panel defined *common assessments* as those that are administered in a routine, consistent manner by a state, district, or school to measure students' academic achievement.[14] These include

- annual statewide accountability tests such as those required by No Child Left Behind;
- commercially produced tests—including interim assessments, benchmark assessments, or early-grade reading assessments—administered at multiple points throughout the school year to provide feedback on student learning;
- end-of-course tests administered across schools or districts; and
- interim tests developed by districts or schools, such as quarterly writing or mathematics prompts, as long as these are administered consistently and routinely to provide information that can be compared across classrooms or schools.

Annual and interim assessments vary considerably in their reliability and level of detail, and no single assessment can tell educators all they need to know to make well-informed instructional decisions. For this reason, the guide emphasizes the use of multiple data sources and suggests ways to use different types of common assessment data to support and inform decision making. The panel recognizes the value of classroom-specific data sources, such as tests or other student work, and the guide provides suggestions for how these data can be used to inform instructional decisions.

The use of data for school management purposes, rewarding teacher performance, and determining appropriate ways to schedule the school day is beyond the scope of this guide. Schools typically collect data on students' attendance, behavior, activities, coursework, and grades, as well as a range of administrative data concerning staffing, scheduling, and financing. Some schools even collect perceptual data, such as information from surveys or focus groups with students, teachers, parents, or community members. Although many of these data have been used to help inform instructional decision making, there is a growing interest among educators and policy advocates in drawing on these data sources to increase operational efficiency inside and outside of the classroom. This guide does not suggest how districts should use these data sources to implement data-informed management practices, but this omission should not be construed as a suggestion that such data are not valuable for decision making.

Status of the research

Overall, the panel believes that the existing research on using data to make

14. The panel recognizes that some schools do not fall under a district umbrella or are not part of a district. For the purposes of this guide, *district* is used to describe schools in partnership, which could be either a school district or a collaborative organization of schools. Technical terms related to assessments, data, and data-based decision making are defined in a glossary at the end of the recommendations.

instructional decisions does not yet provide conclusive evidence of what works to improve student achievement. There are a number of reasons for the lack of compelling evidence. First, rigorous experimental studies of some data-use practices are difficult or infeasible to carry out. For example, it would be impractical to structure a rigorous study investigating the effects of implementing a districtwide data system (recommendation 5) because it is difficult to establish an appropriate comparison that reflects what would have happened in the absence of that system. Second, data-based decision making is closely tied to educational technology. As new technologies are developed, there is often a lag before rigorous research can identify the impacts of those technologies. As a result, there is limited evidence on the effectiveness of the state-of-the-art in data-based decision making. Finally, studies of data-use practices generally look at a bundle of elements, including training teachers on data use, data interpretation, and utilizing the software programs associated with data analysis and storage. Studies typically do not look at individual elements, making it difficult to isolate a specific element's contribution to effective use of data to make instructional decisions designed to improve student achievement.

This guide includes five recommendations that the panel believes are a priority to implement. However, given the status of the research, the panel does not have compelling evidence that these recommendations lead to improved student outcomes. As a result, all of the recommendations are supported by low levels of evidence. While the evidence is low, the recommendations reflect the panel's best advice—informed by experience and research—on how teachers and administrators can use data to make instructional decisions that raise student achievement. In other words, while this panel of experts believes these practices will lead to improved student achievement, the panel cannot point to rigorous research that proves the practices do improve student achievement.

Summary of the recommendations

The recommendations in this guide create a framework for effectively using data to make instructional decisions. This framework should include a data system that incorporates data from various sources, a data team in schools to encourage the use and interpretation of data, collaborative discussion sessions among teachers about data use and student achievement, and instruction for students about how to use their own achievement data to set and monitor educational goals. A central message of this practice guide is that effective data practices are interdependent among the classroom, school, and district levels. Educators should become familiar with all five recommendations and collaborate with other school and district staff to implement the recommendations concurrently, to the extent that state and district resources and capacity allow. However, readers who are interested in implementing data-driven recommendations in the classroom should focus on recommendations 1 and 2. Readers who wish to implement data-driven decision making at the school level should focus on recommendations 3 and 4. Readers who wish to bolster district data systems to support data-driven decision making should focus on recommendation 5. Finally, readers interested in technical information about studies that the panel used to support its recommendations will find such information in Appendix D.

To account for the context of each school and district, this guide offers recommendations that can be adjusted to fit their unique circumstances. Examples in this guide are intended to offer suggestions based on the experiences of schools and the expert opinion of the panel, but they should not be construed as the best or only ways to implement the guide's recommendations. The recommendations, described

Table 2. Recommendations and corresponding levels of evidence

Recommendation	Level of evidence
1. Make data part of an ongoing cycle of instructional improvement	Low
2. Teach students to examine their own data and set learning goals	Low
3. Establish a clear vision for schoolwide data use	Low
4. Provide supports that foster a data-driven culture within the school	Low
5. Develop and maintain a districtwide data system	Low

Source: Authors' compilation based on analysis described in text.

here briefly, also are listed with their levels of evidence in Table 2.

Recommendations 1 and 2 emphasize the use of data to inform classroom-level instructional decisions. Recommendation 1 suggests that teachers use data from multiple sources to set goals, make curricular and instructional choices, and allocate instructional time. It describes the data sources best suited for different types of instructional decisions and suggests that the use of data be part of a cycle of instructional inquiry aimed at ongoing instructional improvement. Building on the use of data to drive classroom-based instructional decisions, recommendation 2 provides guidance about how teachers can instruct students in using their own assessment data to develop personal achievement goals and guide learning. Teachers then can use these goals to better understand factors that may motivate student performance and can adjust their instruction accordingly.

The panel believes that effective data use at the classroom level is more likely to emerge when it is supported by a data-informed school and district culture. Recommendations 3, 4, and 5, therefore, focus on the organizational and technological conditions that support data use. Recommendation 3 suggests that school leaders establish a comprehensive plan for data use that takes into account multiple perspectives. It also emphasizes the need to establish organizational structures and practices that support the implementation of that plan.

The panel believes that effective data use depends on supporting educators who are using and interpreting data. Recommendation 4 offers suggestions about how schools and districts can prepare educators to use data effectively by emphasizing the importance of collaborative data use. These collaboration efforts can create or strengthen shared expectations and common practices regarding data use throughout a school.

Recommendation 5 points out that effective, sustainable data use requires a secure and reliable data-management system at the district level. It provides detailed suggestions about how districts or other educational entities, such as multidistrict collaboratives or charter management organizations, should develop and maintain a high-quality data system.

Checklist for carrying out the recommendations

Recommendation 1. Make data part of an ongoing cycle of instructional improvement

☐ Collect and prepare a variety of data about student learning.

☐ Interpret data and develop hypotheses about how to improve student learning.

☐ Modify instruction to test hypotheses and increase student learning.

Recommendation 2. Teach students to examine their own data and set learning goals

☐ Explain expectations and assessment criteria.

☐ Provide feedback to students that is timely, specific, well formatted, and constructive.

☐ Provide tools that help students learn from feedback.

☐ Use students' data analyses to guide instructional changes.

Recommendation 3. Establish a clear vision for schoolwide data use

☐ Establish a schoolwide data team that sets the tone for ongoing data use.

☐ Define critical teaching and learning concepts.

☐ Develop a written plan that articulates activities, roles, and responsibilities.

☐ Provide ongoing data leadership.

Recommendation 4. Provide supports that foster a data-driven culture within the school

☐ Designate a school-based facilitator who meets with teacher teams to discuss data.

☐ Dedicate structured time for staff collaboration.

☐ Provide targeted professional development regularly.

Recommendation 5. Develop and maintain a districtwide data system

☐ Involve a variety of stakeholders in selecting a data system.

☐ Clearly articulate system requirements relative to user needs.

☐ Determine whether to build or buy the data system.

☐ Plan and stage the implementation of the data system.